Submitted to USFWS
22 June 1988

METHODOLOGICAL ISSUES IN THE ESTIMATION OF TRENDS
IN BIRD POPULATIONS WITH AN EXAMPLE:
THE PINE WARBLER

Frances C. James,[1] Charles E. McCulloch,[2] and Loretta Wolfe[3]

BU-989-M June 1988

ABSTRACT

The Breeding Bird Survey (BBS) is designed to provide estimates of the annual distribution of birds in North America and trends in their abundance. Data taken by volunteers on a single morning each June consist of the numbers of birds by species seen or heard on 50 stops along a 25-mile predetermined route. Approximately 2,000 routes are run annually in the United States and Canada. The analysis of trends across years in this kind of data

[1]Department of Biological Science, Florida State University, Tallahassee, FL 32306-2043.

[2]Biometrics Unit, Department of Plant Breeding and Biometry, Cornell University, Ithaca, NY 14853.

[3]Wolfe Associates, 1706 Semalachee Drive, Tallahassee, FL 32301.

is difficult. Decisions must be made about the quality of the data, how to handle missing data, how to weight the data, and what type of mathematical model to use.

We review the method of analysis developed by Geissler and Noon, which was used in a recent 15-year summary of BBS data, and we compare it with a method developed by Mountford for the analysis of censuses made by the spot-map method of censusing in England. Then we propose a method that combines features of each of the above methods but also has several features that are new. We use the number of stops on which a species was recorded as a criterion of abundance. We define complete routes as those for which there are at least 10 years of reliable data, distributed so that in the 22-year period 1966-1987 each 5- or 6-year period is represented. For the Pine Warbler (Dendroica pinus) in the Central Southern Region of the United States, 201 routes meet these criteria. For the model to fit, we use LOWESS, a method of robust nonparametric regression, which smooths data across years within each route. Then we pool smoothed route values by physiographic strata within states for 37 stratum-within-state units. Trends for states, for strata, or for the region as a whole are derived from aggregated values, weighted by the proportion of the larger area that is occupied by each unit. We explore the value of smoothing the data at each step of aggregation. Our method allows the construction of graphs expressed in an original unit of abundance (stops per route). We think that results based on the proposed method should be more useful to managers than is information restricted to linear trends in data, as with the Geissler and Noon method, or to unsmoothed data, as in the Mountford method. Two of the advantages of the proposed method are, first, that abundance data are fit directly on a route-by-route basis--so no abundance weighting or log transformations are needed--and, second, that the nonparametric smoother, LOWESS, allows flexibility in the degree of smoothing used.

INTRODUCTION

The Breeding Bird Survey (BBS) is the most ambitious attempt to monitor populations of animals in nature that has ever been conducted. Experienced volunteers follow predetermined routes for 25 miles, stopping each half-mile to record the birds seen or heard in a three-minute period. The results are compiled annually by the U.S. Fish and Wildlife Service and summarized in maps and graphs showing indices of the abundances of the species and trends in their populations through time. The major summary of this project is by Robbins et al. (1986) for the 15-year period 1965-1979. In this publication (and in Droege, this volume) the authors discuss the important difficulties and choices that must be made before analysis of the data. Here we review the subjects of data quality, what to do about missing data, how to aggregate data by routes into trends for larger areas, how to weight the data, and how to choose a model to fit. We compare the method of Geissler and Noon (1981), which was used in the 15-year summary, with a method developed by Mountford (1982, 1985) for the analysis of censuses made on plots in England. Then we propose a new method that is based on first smoothing data for each route by a locally weighted regression (LOWESS, Cleveland 1979, 1981; Chambers 1983) and then aggregating the resulting values by physiographic strata within states. To illustrate this method, we analyze data for the Pine Warbler (*Dendroica* *pinus*) in the central southern and southeastern region of the United States for the 22-year period 1966-1987. An outline of the recommended procedure is given in the appendix.

STATISTICAL ISSUES

In this section we review methodological issues that influence the effectiveness of statistical approaches to the analysis of the BBS data. The first issue is the quality of the basic data. Without high-quality data, no statistical method will perform satisfactorily. After the basic issue of data quality, there is the question of what data are available. A consequence of such a large, volunteer effort as the BBS is that for many years there are routes with no data; i.e., there are missing data. Subsequent to decisions on questions of data quality and availability come issues of aggregation and amount and type of weighting. These issues require decisions about the geographic and temporal scale on which the analysis is to be performed and about how geographic or temporal groupings will be combined to get overall analyses. Finally, we discuss appropriate types of models to fit to the data in order to summarize the results.

Data Quality

Without high-quality data, no statistical method will give reliable analyses. Because the BBS depends on volunteers, whose work can only be minimally supervised, data quality is a big concern. We will not go into detail here, as there are thorough discussions both in this volume and elsewhere (Robbins et al. 1986). However, one pitfall to keep in mind with data of this sort is the belief that the sheer volume of the data will somehow make things better, even if the basic data quality is poor. To paraphrase Freedman et al. (1978, p. 303), "If a measurement procedure leads to biased estimates, taking a large sample doesn't help. It just repeats the basic mistake on a larger scale." Thus, care must be expended to improve and monitor the data quality, even if data quantity must be sacrificed.

Missing Data

One of the most difficult facets of developing appropriate statistical methodology for the BBS is the large percentage of missing data. If there were an association between abundance measurements and the pattern of missing data, then ignoring the missing data could lead to erroneous conclusions. For example, if there were more missing data in later years of a survey from the routes of higher abundance, then the population trend would be underestimated or estimated to be negative when it was not. Consider the example given in Table 1. There is no population trend at any location. Yet, because there is an association between the missing data and abundance (there are more missing data at the sites of higher abundance), there appears to be a decline when only the averages are considered. A simple average of the routes with data present puts more weight on the routes with smaller abundances in the later years.

Table 1. An association between missing data and abundance can falsely indicate a population decline when only the averages are considered.

	Year				
Route	1	2	3	4	5
1	32	32	.	.	.
2	20	20	20	.	.
3	14	14	14	14	.
4	2	2	2	2	2
Average	17	17	12	8	2

It is possible to use the data that are present to estimate the pattern for years for which data is missing. Both the Giessler-Noon method and the Mountford method do so by fitting a model that incorporates route effects. In Table 1, if the missing data are first estimated within route and are then averaged, the results show no trend.

Aggregation

The level of aggregation is also a very important decision for the statistical analysis. Premature aggregation on a geographic level hides trends at a route level. For the data in Table 1, it is the premature aggregation of the data across routes that leads to the misleading decline in the averages. The aggregation of data temporally affects what trend is being estimated. It is easily possible for a population trend to be positive for one group of years and negative for another. A successful method of analysis should allow for the relatively automatic discovery of different trends for different choices of the years to be analyzed.

Weighting

If the analysis is performed at a level of aggregation that does not incorporate the entire data set, then the question of weighting needs to be addressed. Will groups in the analysis receive equal weighting? If not, by what criteria should the estimates be weighted? Common weighting factors are abundance, area represented, and statistical precision. The choice of whether to weight by each of these criteria and how to do so is problematic. We consider each in turn. For concreteness we assume that the analysis is performed at the route level, as in the Geissler-Noon method.

If routes are representative of the areas in which they are located, then it is a good idea to weight by the area represented by each route. This would generally be the case for the BBS but not, for example, for the Colonial Bird Survey. There, each colony represents only the abundance at that colony, not the abundance in a surrounding area. Hence, it would not make sense to use area weighting.

If trends in the overall population are dessired, then weighting by abundance should be performed. If an area with a very large population size is declining and an area with small population size is increasing, then the overall trend reported should be a decline. For an illustration, see the example in Table 2, in which one route shows a halving of the population size each year and the other a doubling. The overall trend is a severe decline. This decline would only be evident in an overall analysis that weighted by abundance. On the other hand, there are situations in which abundance weighting should not be used. For example, if the proportion of routes that are declining is the information sought, then abundance weighting would not be appropriate.

Table 2. An overall decline in a population can be adequately illustrated by averages, but it would be masked if unweighted averages of the trend were calculated.

	Year				Trend (Annual
Route	1	2	3	4	proportional change)
1	192	96	48	24	.50
2	2	4	8	16	2.00
Average	97	50	28	15	1.25

Weighting by statistical precision should be used. However, the proper method is not at all clear. It is not sufficient to weight by the within-route estimated variance. For an overall analysis, the observations within a route must necessarily be considered correlated, and therefore the weighting scheme must account for the amount of dependence within routes. Such a scheme would require modeling of, or estimation of, the dependence, to get estimates of the between- and within-route variability of the trend estimates.

Model to Fit

Finally, there is the question of the statistical model to fit, the one that will effectively summarize the trends evident in the data. The model that is fit to the data determines how the trends are estimated and, among other things, determines the "smoothness" of the trend estimates. It does not have to be the "true" model, or even necessarily a realistic model to summarize a trend effectively. If a model function is fit that requires the trend to be a very simple function, then the trend will necessarily be "smooth." If there is a large amount of flexibility in the estimated trend, then the trend may be very "bumpy."

The type of model used will affect the efficiency of estimates. A fully efficient method should exploit the correlations in the data within a route. Thus, the model should be sensitive to the corelation structure in the data.

A COMPARISON OF THE GEISSLER-NOON AND MOUNTFORD MODELS

In this section, we compare the methods of Geissler-Noon and Mountford with respect to the criteria identified in the previous section. We discuss how each handles missing data, aggregation level, weighting, and the model that is fit.

Both the Geissler-Noon and Mountford methods have mechanisms for handling missing data. In the Geissler-Noon methodology, which analyzes the data on a log scale, a separate regression line is fit for each route, allowing each route to have its own intercept and slope. The estimated slopes are then combined by means of a weighted average. This method avoids the problem (see Table 1) of a route's having no weight for years in which the data are missing. In a somewhat similar fashion, the model for the Mountford analysis has site effects. This feature acts as a type of correction for missing data. Thus, each of these methods of analysis handles missing data in a suitable manner.

In terms of aggregation, the Geissler-Noon method performs its analysis on a route-by-route basis and then aggregates to larger geographic levels. Temporally, the Geissler-Noon analysis uses the entire set of years to find a single estimate of trend. Of course, a subset of the years can be reanalyzed to search for trends over shorter time periods. The Mountford analysis does not address either geographic or temporal aggregation. In the analysis, separate effects are estimated for each year, so temporal aggregation is not an advantage. In essence, the Mountford analysis measures yearly fluctuations, not trends. Geissler (1985) has suggested a supplement to the trend estimate for estimating annual indices. This feature is essentially the same as the Mountford model (detailed below) and estimates separate effects for each year. Sauer (this volume?) has suggested using a method of residual analysis to modify the Geissler-Noon method. This method also gives yearly fluctuations as a supplement to the Geissler-Noon trend analysis.

The Geissler-Noon analysis uses weights equal to the product of area, abundance, and the inverse of the variance of the slope estimate. As mentioned earlier, weighting by area and abundance is usually the correct weighting scheme. However, weighting by the inverse of the variance of the slope estimate, calculated on the assumption that the observations within a route are independent, is not an optimal weighting. An extreme example serves to illustrate why. If the data from a route are nearly perfectly fit by the Geissler-Noon model, then that route would have nearly infinite weight and would dominate the analysis, swamping all the data from the other routes. Because there is substantial route-to-route variation in the BBS data, a route with a good within-route fit should not be allowed to overshadow other routes greatly. The Mountford analysis does not use any type of weighting. However, the method could be adapted to include any or all of the above-mentioned weighting schemes.

In many ways, the models fit by the Geissler-Noon analysis and the Mountford analysis are similar. The Geissler-Noon model is a linear regression through time on the log scale. The Mountford analysis uses a similar multiplicative model. More precisely, if we let C_{yr} denote the count for year y and route r, then the Geissler-Noon model, ignoring observer effects, is given by:

$$\log (C_{yr} + .5) = \log (a_r) + y * \log (b_r) + \text{error}, \qquad (1)$$

and the Mountford model is given by:

$$C_{yr} = a_r * b_y + \text{error}. \qquad (2)$$

From this equation, we see that the Mountford analysis uses an additive error on the original scale, whereas the Geissler-Noon analysis uses an additive error on the log scale. If we consider only the mean values and ignore the constant of .5 that is added to correct for zero counts, then the models can be written as

$$\log (C_{yr}) = \log (a_r) + y * \log (b_r), \tag{3}$$

for Geissler-Noon and

$$\log (C_{yr}) = \log (a_r) + \log (b_y), \tag{4}$$

for the Mountford analysis. From these equations, we can see that the two analyses are similar, fitting additive models on the log scale, with separate constants for each route. They are different, however, in that the Geissler-Noon analysis fits separate slopes for each route but the same change from year to year, whereas the Mountford method fits separate effects for each year but assumes that they are common to all the routes. This different treatment of the year effects shows how the methods produce estimates of changes from year to year. The Geissler-Noon analysis summarizes the information by estimating the linear trend on the log scale for the entire period of time considered in the analysis. This method produces as "smooth" a trend estimate as is possible. The Mountford method is at the other extreme, fitting each year with a separate effect, allowing the ultimate in flexibility. It will tend to yield very "bumpy" estimates of the trend through time.

The Geissler-Noon and Mountford analyses handle variances and covariances in the data in extremely different manners. The Geissler-Noon method ignores the correlation within a route when fitting separate slope estimates. It uses the between-route variability to estimate standard errors of the trend estimates. Because the trends are estimated directly, confidence intervals on the mean abundance cannot be calculated. Hence the diagrams by Robbins et al. (1986) that give confidence intervals are for the trend estimtes only. They cannot be interpreted as confidence bands for the mean abundance line as a function of time. The Mountford method estimates the correlations in the yearly ratio estimates to obtain slightly more efficient estimators. Either method is suitable for the accommodation of the correlations in the data and the calculation of standard errors.

How might these two methods be improved? They both handle missing data suitably, and the Geissler-Noon method addresses aggregation well by building from a route-level analysis to represent larger geographic regions. We can suggest improvements in the other two areas: weighting and model to be fit. If the data are fit directly on a route-by-route basis, then no abundance weighting need be done. Fitting the data directly also avoids the complications involved with logarithms and the corrections needed for zero counts. If the major component of variation is route-to-route variation, then routes representing equal areas should be weighted approximately equally, and the within-route variationshould essentially be ignored. Though still not optimal, this is the weighting we have chosen. With respect to the model to be fit, we suggest one that is an intermediate between Geissler-Noon and Mountford. Mountford's method does not use any of the information from year to year to try to remove extraneous variation that is not part of a trend. On the other hand, the Geissler-Noon method produces only a single estimate for the entire period of time considered and hence does not help to discover trends over shorter time periods. We suggest the use of a method that allows the estimates to be smoothed somewhat but not forced to be straight lines on the log scale.

Because each route is run only once each year, some of the between-year variation is undoubtedly attributable to factors other than long-term trends in the bird population. Trends should be more evident if some of this extraneous variation is removed by smoothing. We selected LOWESS (locally weighted scatterplot smoother) (Cleveland 1979, 1981; Chambers 1983) for two reasons. First, it allows us to judge trends without having to select in advance the years for which the trend is to be calculated. For example, a trend that began in the middle of the 15-year period of the Robbins et al. (1986) analysis would be more sensitively detected by our method that by a linear trend analysis. The second advantage of LOWESS is that it is a conditional smoother; that is, it entails no assumption that the values on the x axis (years) are equally spaced. The degree of smoothing with LOWESS is optional. If f is set to 1.0, the smoothing is complete, and the result is nearly a straight line. If f is set to .5, half of the data are scanned to calculate each fitted value, and very general patterns can be expressed.

In Figure 1, a single route from the BBS Scissor-tailed Flycatcher data (No. 7007) is fit with LOWESS. It illustrates the effect of changing the smoothing parameter, f. Small values of f (near zero) allow many changes in the fitted line; large values of f (near 1) lead to nearly linear fitted lines. Table 3 gives hypothetical data and shows how the variances and confidence intervals can be calculated for the LOWESS lines using route-to-route variability. The data consist of six routes in one stratum-within-state and seven routes in the other stratum-within-state. The first stratum-within-state is assumed to be 100 units in area, and the other is assumed to be 50 units. Because these calculations are simply weighted averages, there is no need for bootstrapping.

AN EXAMPLE OF THE PROPOSED METHOD: THE PINE WARBLER

To explore the usefulness of the proposed method for the examination of trends in a particular bird population, we present an analysis of data for the Pine Warbler, a common breeding bird of pine woods from southern Canada to the Gulf of Mexico. We consider only the nine-state area of the southern central and southeastern United States (called the central southern region) (Fig. 2, Table 4).

Data Quality, Missing Data, Aggregation

The BBS data for each route in each year have been rated by the staff of the Fish and Wildlife Service on the basis of weather and the reliability of observers. We have used only type 1 data, the type they judged to be most reliable. For the 22-year period 1966-1987, there were 399 routes in the Central Southern Region for which there is some type 1 data. However, many routes were not run in all 22 years, and the data for some were judged to be

Table 3. Calculation of mean and variance in LOWESS estimated values by route into stratum-within-state units and aggregation of these value to get average state values and their variances, without smoothing above the route level.

Stratum-within-state unit 1 Area = 100	Route	LOWESS values by year (stops per route) 1	2	3	4	5
	1	0	0	0	1	1
	2	1	3	4	5	6
	3	2	4	6	7	7
	4	1	1	1	1	1
	5	2	3	4	2	2
	6	4	5	4	4	4
Stratum-within-state 1 mean		1.67	2.67	3.17	3.33	3.50
Estimated variance of mean (s^2/n)		.31	.58	.83	.98	1.12

Stratum-within-state unit 2 Area = 50	Route	LOWESS values by year (stops per route) 1	2	3	4	5
	1	1	3	4	5	1
	2	2	4	4	4	3
	3	5	5	4	4	3
	4	6	6	7	6	6
	5	1	2	0	1	2
	6	1	1	1	1	1
	7	3	3	5	3	3
Stratum-within-state 2 mean		2.71	3.43	3.57	3.43	2.71
Estimated variance of mean (s^2/n)		.71	.49	.94	.60	.48

Table 3 (continued).

Weighted average = $\Sigma w_i L_i = \dfrac{100L_1 + 50L_2}{150}$

Variance = $\Sigma w_i{}^2 s_i{}^2 = (100/150)^2\ (s_1{}^2/n_1) + (50/150)^2\ (s_2{}^2/n_2)$

	Year				
	1	2	3	4	5
State weighted average	2.02	2.92	3.30	2.37	3.24
Variance of weighted average	.21	.30	.46	.49	.54

For approximate confidence intervals, use $\pm\ 2\sqrt{\text{variance}}$

---

less reliable for certain years. For the 399 routes on 22 years, approximately 4,700 of the possible 8,778 (22 x 399) records are available. Our criterion for inclusion of a route with missing data was that at least 10 years be represented and that there be at least one year of data in each of four periods: the first six years, each of the next five-year periods, and the last six years. The 201 routes that met these criteria were deemed sufficiently "complete" to be included in our analysis (Fig. 2). Note that the selection of routes was not affected by the presence of Pine Warblers but only by the reliability of the survey and the coverage of the period of interest.

Because of insufficient data, we were not able to include Louisiana or South Carolina in our analysis. In each of these states, there were long series of years in which no surveys were conducted. Also, we excluded stratum/state units for which there were fewer than three routes (stratum 1 in Florida, strata 3 and 13 in Georgia, and stratum 4 in North Carolina). Of

Table 4. The proportion of the area of each physiographic stratum in each state, the proportion of each state in the Central Southern Region, and the number of routes with reasonably complete and reliable data by state.

	AL	AR	FL	GA	LA	MS	NC	SC	TN	Total
Area (km^2/1000)	81.0	83.8	87.2	116.2	72.5	76.3	78.7	47.8	66.7	710.2
Proportion of region	.114	.118	.123	.164	.102	.107	.111	.067	.094	1.00
No. of complete routes	37	28	32	18	12	14	11	10	39	201
No. of strata	5	3	4	5	4	3	4	3	6	37
Proportions of strata within states										
1 Subtropical			.130							
2 Floridian			.454							
3 Lower Coastal Plain	.028		.327	.113	.112	.037	.264	.267		
4 Upper Coastal Plain	.560	.238	.088	.378	.340	.787	.199	.372	.212	
5 Mississippi Alluvial Plain		.289			.304	.177			.024	
6 East Texas Prairies					.245					
11 Southern Piedmont	.078			.239			.365	.360		
13 Ridge and Valley	.269			.239					.183	
14 Highland Rim	.066								.409	
19 Ozark Ouachita Plateau		.473								
21 Cumberland Plateau									.111	
23 Blue Ridge Mountains				.030			.172		.061	
	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

course we included all 39 complete routes in Tennessee, even though the numbers of Pine Warblers were very low (Table 4).

We have not addressed the issue of estimation of missing values. In one case (estimation of the state line for Georgia for 1969 and 1970) we used an average to get a missing value. Furthermore, we analyzed only 201 of the 399 routes for which there were type 1 data. By setting a criterion for the estimation of missing values, one could use more of the data. Whether or not this procedure would change the result would have to be determined.

The Choice of an Abundance Criterion: Stops per Route

There are two logical choices for a criterion of abundance for Breeding Bird Survey data: the total number of individuals of the species recorded on the 50 stops of one survey and the number of stops out of the total of 50 at which the species was recorded. Robbins et al. (1986) used the former; Cox (1987) used the latter. Of course the variable of most interest is the number of birds present, but there are several reasons why stops per route might be a more reliable indicator of abundance than the actual count of individual birds. First, it is less difficult for an observer to determine whether or not a species is present at a particular stop than it is to determine how many individuals of each species are present. Second, there is probably a greater difference between observers in judging the number of individual birds than there is in determination of whether the species is present (Bart and Schoultz 1984; S. Droege, pers. comm.). Also, analysis by stops per route would be simpler than analysis by individuals. For these reasons, a high positive correlation between these two criteria would be justification for using stops per route on which the species was recorded as a measure of abundance.

We examined data for 18 routes in the Lower Coastal Plain (stratum 3) in northern Florida. Data for all 18 routes were pooled, and average stops per route on which Pine Warblers were recorded was plotted along with the average number of individual Pine Warblers that were seen or heard per route (Fig. 3). The two variables are highly correlated. When their standard deviations by year are compared, the variation around these average values is less for stops per route. All of these factors supported our decision to use stops per route as our abundance criterion.

Model to Fit: The Choice of a Smoother

A plot of the original data values and the smoothed LOWESS line with f = .5 for three routes in the Upper Coastal Plain stratum (stratum 4) in southern Arkansas illustrates the value of some smoothing. The trends are clear in the original data, but they are emphasized in the LOWESS lines (Fig. 4a). Smoothed lines for all nine routes in this stratum in Arkansas allow visual comparison of substantially more data, and a second smoothing of averages of all of these values expresses the general trend for the stratum-within-state unit (Fig. 4b). A comparison of this last line with the original values shows that the generally increasing S-shaped trend is very difficult to see in the morass of the original lines (Fig. 4c). Note, also, that the low values for the summer of 1978, which followed two extremely cold winters, have been obscured by the doubly smoothed stratum-within-state line. This example demonstrates the importance of selecting the degree of smoothing on the basis of the generality of the particular question of interest. If the question is to find the effects of particular events on bird populations, smoothing may be unwise. Expression of the data at more than one level of smoothing can be very helpful, but even mild smoothing can shift the time of the apparent occurrence of a change in the size of a population.

Aggregation and Smoothing

Because it is important to minimize bias attributable to differences among observers and to other site effects, we analyze records for each route separately. Then, as in the method of Geissler and Noon (1981), we average route values into units that are first physiographic strata within states. We obtain state or stratum values by averaging stratum-within-state values and weighting by proportional areas. A stratum that has a strong increase in Pine Warblers will not affect a state trend much if that stratum occupies only a small area of the state. See for example stratum 4 in Florida (Table 4, Fig. 5d). We constructed state and regional trend lines using weighting proportional to area (Table 4) as suggested by Geissler and Noon (1981), but we have not weighted the data by precision as they do. Finally, we obtain regional values by averaging state or stratum values and again weighting by proportional area. With this aggregative method, trends at various spatial scales can be compared graphically. In each state there are three to six stratum-within-state units (Table 4, Fig. 2).

The appearance of trend lines is affected by both the level of smoothing and whether smoothing is applied at various levels of aggregation. If data are smoothed at the route level and then aggregated by calculation of means to get trends for stratum-within-state units and state trends, each weighted by proportional area, the trends are somewhat less clear than they are if the smoothed values by route are smoothed again at the higher levels. For Arkansas, Florida, and North Carolina, the successively smoothed lines are given in Fig. 5b, d, and f, and the lines based on only initial smoothing by routes and then aggregation by weighted means are given in Fig. 5a, c, and e. In each case the state line is the solid line.

A trend line for the Central Southern Region can be constructed from values for states (Fig. 5), if they are weighted by the proportion of the region occupied by each state. Again the sensitivity of the trend to variation will depend on the amount of smoothing. If smoothing is applied only to the route data and means are used to aggregate successively larger areas, the regional trend (Fig. 6a) can show the effects of individual years. Note, for instance, the decline in 1981 in Alabama, Georgia, and North Carolina, an effect that is particularly evident in the Southern Piedmont (stratum 11) of North Carolina. To discover whether this is a coincidence or a real decline would take further work. Such phenomena are not apparent when the data are smoothed at each stage of the aggregation (Fig. 6b), but such a plot gives a nearly linear regional trend against a background of less smoothed state lines. The generality of the trend is accentuated at larger spatial scales.

The analysis based on data that seemed hopelessly heterogeneous at the beginning has been reorganized and simplified. All states show clearly increasing populations of Pine Warblers.

In Fig. 6d, it is clear that in the Central Southern region the Pine Warbler has been especially abundant in the Lower Coastal Plain (stratum 3) throughout the 22-year period. Note that the physiographic region in which the population has shown the greatest increase is the Southern Piedmont (stratum 11, mainly Georgia and the Carolinas). These increases plus increases in the Upper Coastal Plain (stratum 4) account for most of the large increase in the population as a whole (Fig. 6d). In Alabama, Arkansas, Florida, and Mississippi, there were slight declines in the late 1970's and early 1980's.

A widespread decline in the number of Pine Warblers occurred after the severe winters of 1976-77 and 1977-78 (see Fig. 3 for northern Florida, Fig. 4c for southwestern Arkansas, Fig. 6a for the entire region). Another apparent decline occurred in 1981 in North Carolina, Georgia, and Alabama (Fig. 6a), especially in the Southern Piedmont stratum (Figs. 6c, d, 7). This decline is associated with a drought that began in May 1980 and continued through an exceptionally cold winter and into the very hot summer of 1981. Were normal numbers of Pine Warblers present but just not singing on these very hot, dry June days? The numbers returned to normal values in 1982, and general increases have continued through the 1980's. Increases in the abundance of Pine Warblers are exceptionally large in the Lower Coastal Plain and in the Southern Piedmont. The population may be responding to changes in the management of pines for saw timber and pole timber in these areas. The highest density of breeding pairs recorded on a Breeding Bird Census in the southeastern coastal states was 30 pairs per 40 ha in sapling loblolly pine (*Pinus taeda*) and shortleaf pine (*Pinus echinata*) pole timber (Hamel et al. 1982).

The Pine Warbler is a member of the genus *Dendroica*. Its congeners are mostly neotropical migrants. The genus is of particular concern because census data taken independently of the BBS suggest that populations of the neotropical migrants are declining (Hall 1984). Published analyses of BBS data have not shown declines for this group (Robbins et al. 1986; Cox 1987). Biologists and conservationists should give serious attention to why there is a disparity between the results of different types of data. We selected the Pine Warbler for study with the idea that future comparisons with its congeners should be instructive.

LITERATURE CITED

Bart, J., and J. Schoultz. 1984. Reliability of singing bird surveys: changes in observer efficiency with avian density. Auk 101:307-318.

Chambers, J.M., W.S. Cleveland, W.S. Kleiner, and P.A. Tukey. 1983. Graphical Methods for Data Analysis. Duxbury Press, Boston, Mass.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Statist. Assoc. 74:829-836.

Cleveland, W.S. 1981. LOWESS: A program for smoothing scatterplots by robust locally weighted regression. Am. Statistician 35:54.

Cox, J. 1987. The Breeding Bird Survey in Florida: 1969-1983. Florida Field Naturalist 15:29-44.

Freedman, D., R. Pisani, and R. Purves. 1978. Statistics. W. W. Norton and Co., New York.

Geissler, P.H. 1985. Estimation of animal population trends and annual indices from a survey of call-counts and other indicators. Proc. Am. Statist. Assoc., Section on Survey Research Methods, 472-477.

Geissler, P.H., and B.R. Noon. 1981. Estimates of avian population trends from the North American Breeding Bird Survey, Studies in Avian Biology 6:42-51.

Hall, G.A. 1984. Population decline of neotropical migrants in an Appalachian forest. Am. Birds 38:14-18.

Hamel, P.B., H.E. LeGrant, Jr., M.R. Lennartz, and S.A. Gauthreaux, Jr. 1982. Bird-habitat relationships in southeastern forest lands. USDA, Forest Service, General Technical Report SE22.

Mountford, M.D. 1982. Estimation of population fluctuations with application to the Common Bird Census. Appl. Statist. 31:135-143.

Mountford, M.D. 1985. An index of population change with application to the Common Bird Census. Pp. 121-132 in D. Brillinger, S. Fienberg, J. Gani, J. Hartigan, and K. Krickenberg (eds.). Statistics in Ornithology. Springer-Verlag, New York.

Robbins, C.S., D. Bystrak, and P.H. Geissler. 1986. The Breeding Bird Survey: its first fifteen years, 1965-1979. USFWS Resource Publication 157.

Tate, J., Jr., and D.J. Tate. 1982. The blue list for 1982. Am. Birds 36:126-135.

Wilcove, D.S., and J.W. Terborgh. 1984. Patterns of population decline in birds. Am. Birds 38:4-13.

FIGURE LEGENDS

Figure 1. Population trend for the Scissor-tailed Flycatcher on one route showing successively more smoothing in a, b, and c.

Figure 2. The Central Southern Region of the United States showing the locations of 201 Breeding Bird Survey Routes for which there are reliable data for a reasonable span of the period 1966-1987 (see text). The physiographic strata (Robbins et al. 1986) are shown by dotted lines and are identified in Table 4.

Figure 3. The average number of stops per 50-stop route on which the Pine Warbler was recorded, and its standard deviation, plotted along with the average total number of individual Pine Warblers seen or heard and its standard deviation. The data are mean values for 18 routes in stratum 3 (Lower Coastal Plain) in Florida.

Figure 4. a. The abundance of Pine Warblers on three of the nine routes in stratum 4 (Upper Coastal Plain) in southern Arkansas. Trends smoothed by LOWESS (f = .5) are given by heavy lines for each route. b. LOWESS-smoothed lines for the nine routes in this stratum-within-state unit and a solid line for resmoothed values for the general trend in this unit. c. The same solid line as in Fig. 4b plotted with the original unsmoothed values for the nine routes.

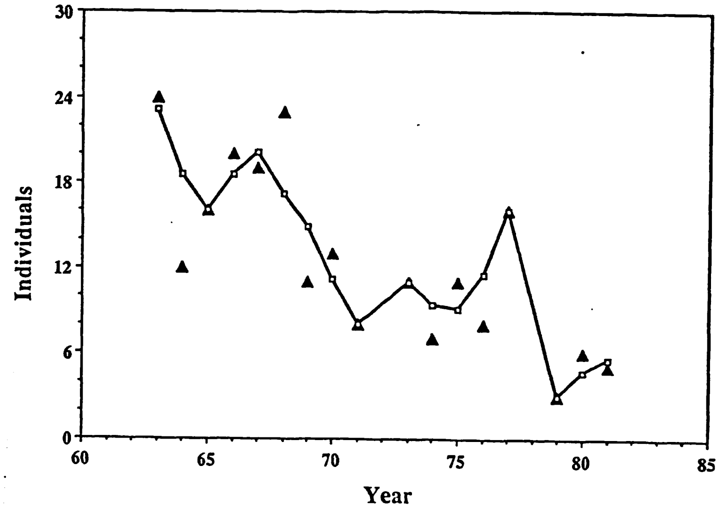
Figure 5. Stratum-within-state trends (dashed lines) and state trends (solid lines) for Arkanasa, Florida, and North Carolina. Lines in Figs. 5a, c, and e were smoothed only at the route level. Lines in Figs. b, d, and f were smoothed at the route, stratum-within-state, and state levels.

Figure 6. State, stratum, and regional trends. Figure 6a gives state lines smoothed only at the route level. In Figure 6b, smoothing was applied at the route, stratum-within-state, state, and regional levels. In Figure 6c, stratum lines were smoothed only at the route level. In Fig. 6d, smoothing was applied at the route, stratum-within-state, stratum, and regional levels.

Figure 7. Trend in the population of the Pine Warbler in the Central Southern United States for the 22-year period 1966-1987 estimated by either state or stratum with two levels of smoothing.

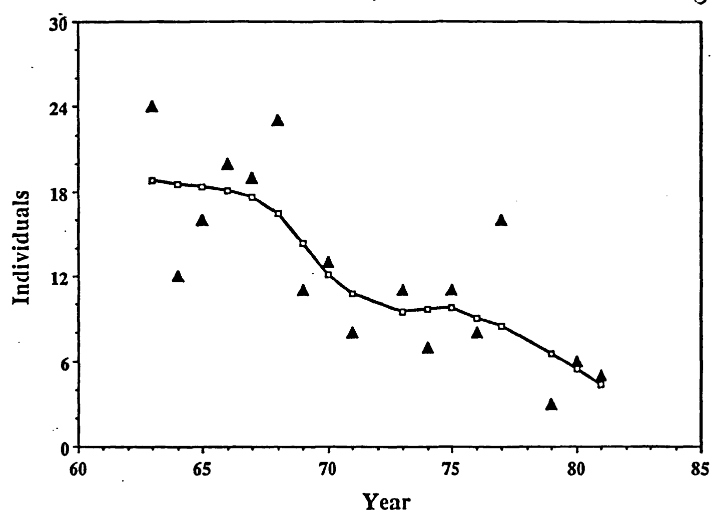
Lowess Smoothed, F = .25

a



Lowess Smoothed, F = .50

b



Lowess Smoothed, F = 1.0

c

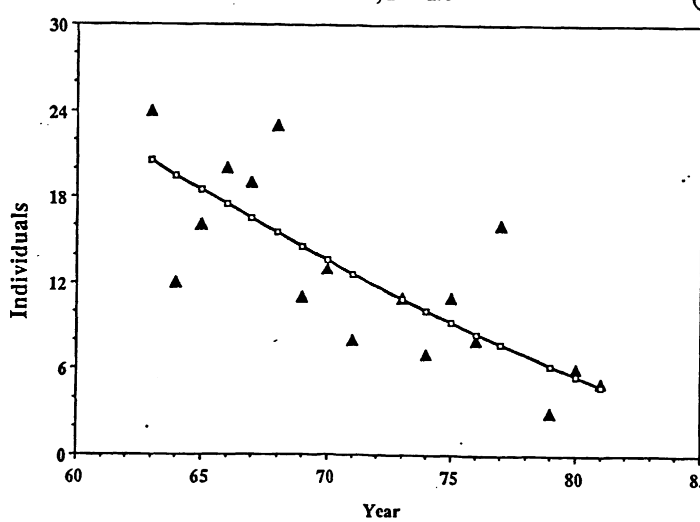
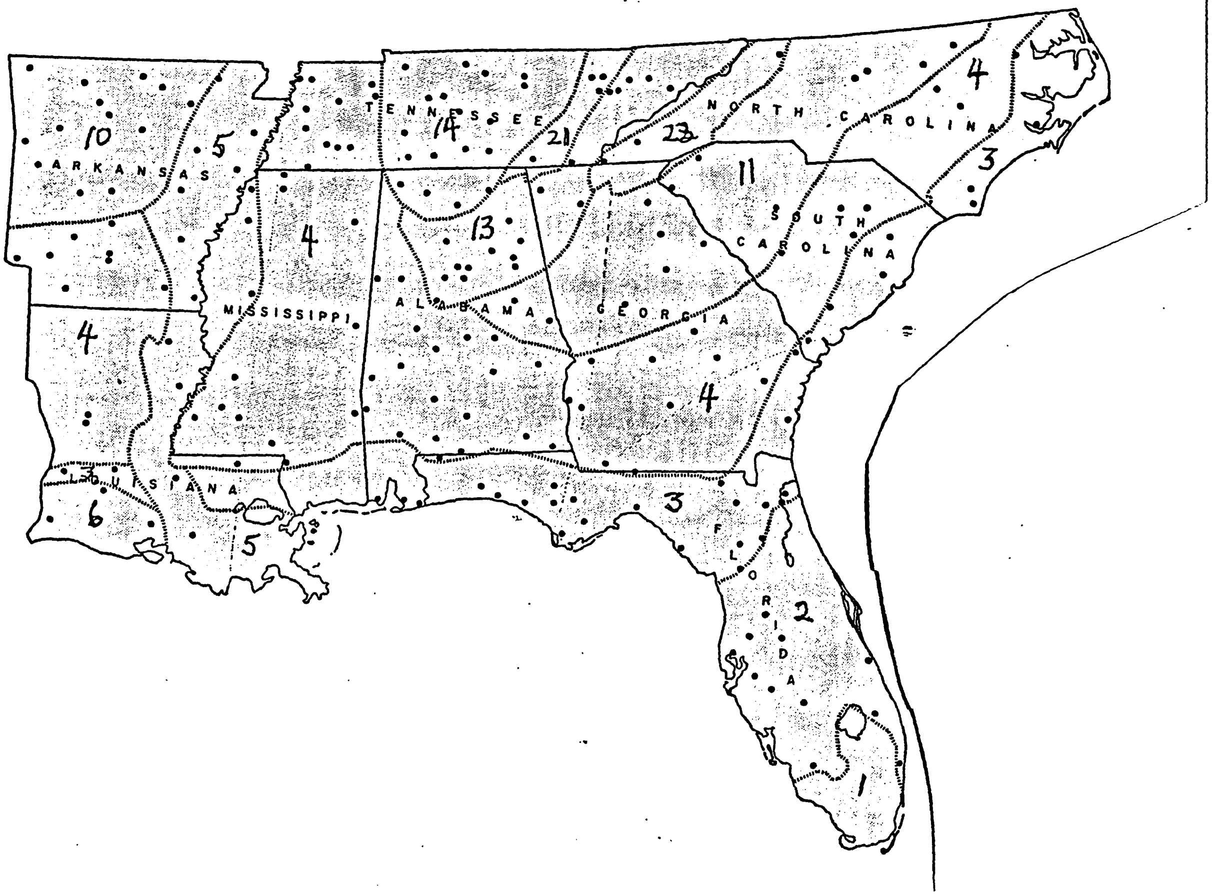


Fig. 2



ARKANSAS
TENNESSEE
NORTH CAROLINA
SOUTH CAROLINA
MISSISSIPPI
ALABAMA
GEORGIA
LOUISIANA
FLORIDA
10
5
14
21
23
4
3
11
4
13
4
4
6
5
3
2
1

Fig. 3

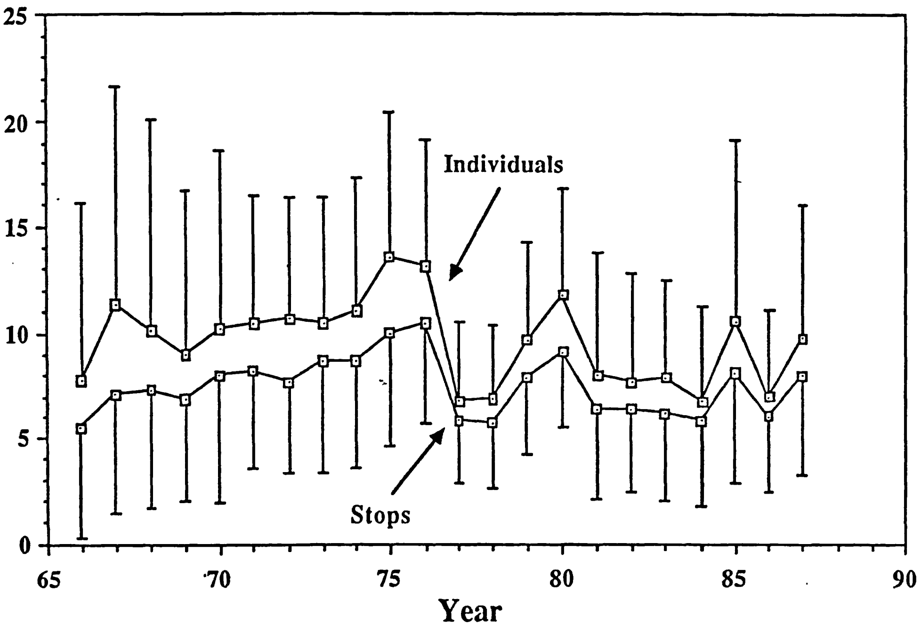
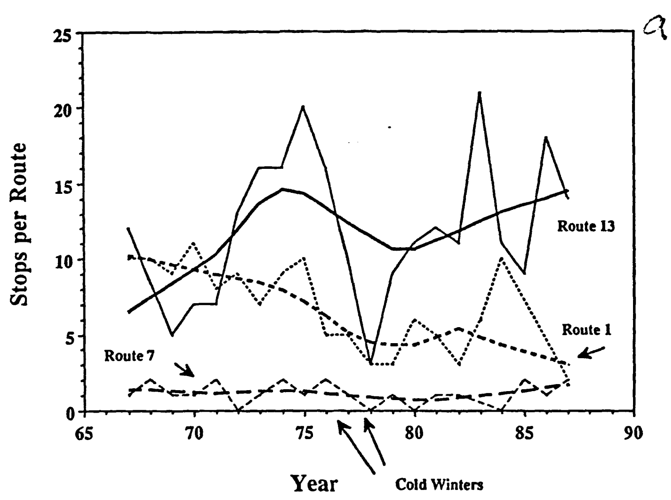
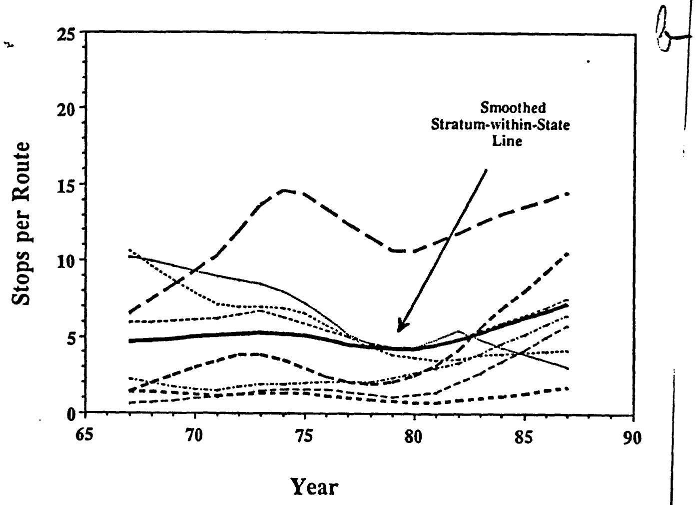


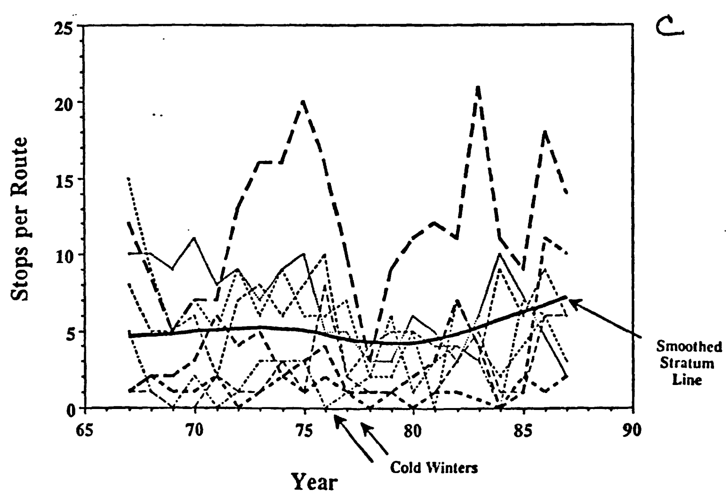
Fig. I



a
Stops per Route
Year
Route 13
Route 1
Route 7
Cold Winters
25
20
15
10
5
0
65
70
75
80
85
90

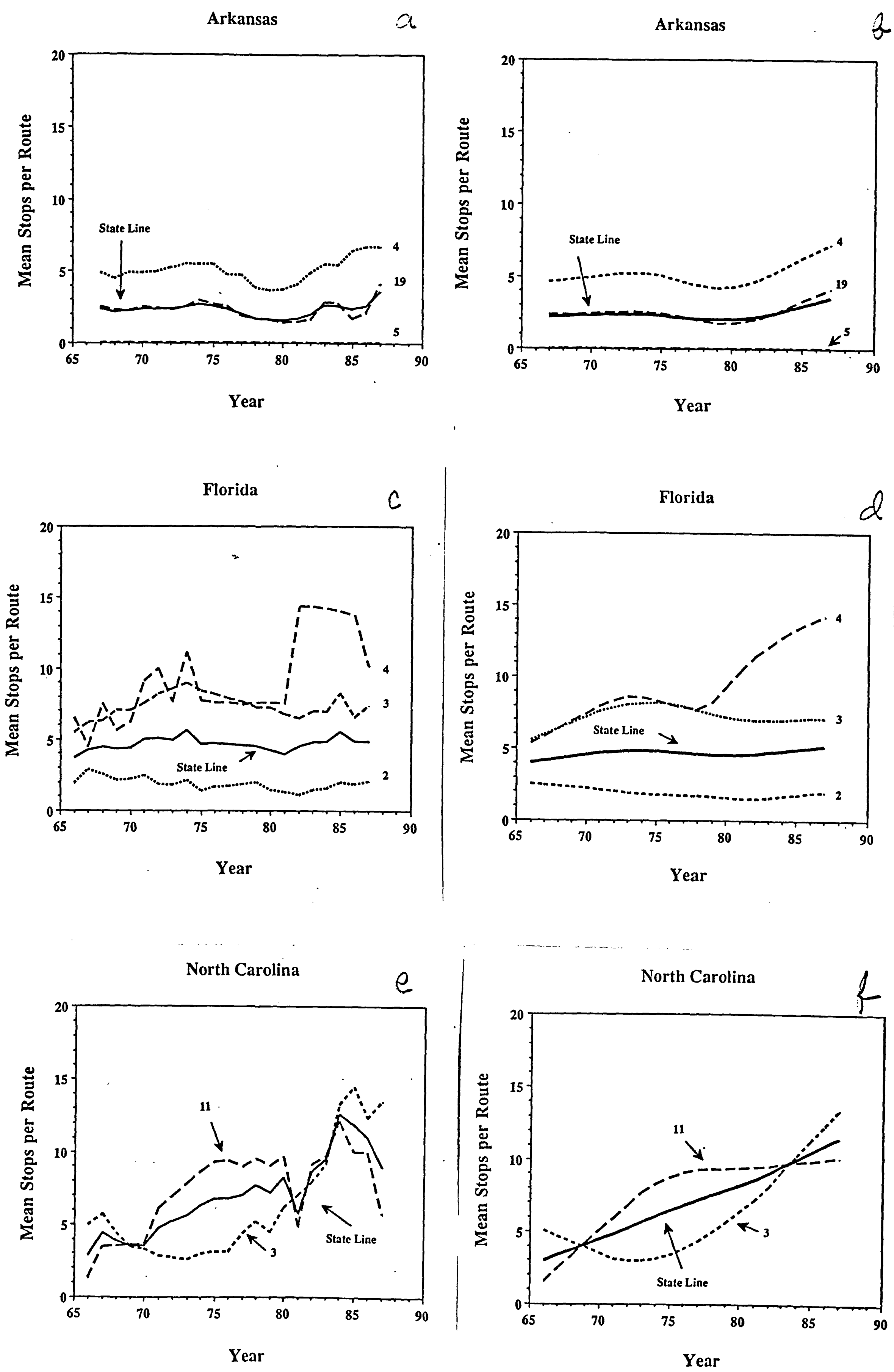


b
Stops per Route
Year
Smoothed
Stratum-within-State
Line
25
20
15
10
5
0
65
70
75
80
85
90



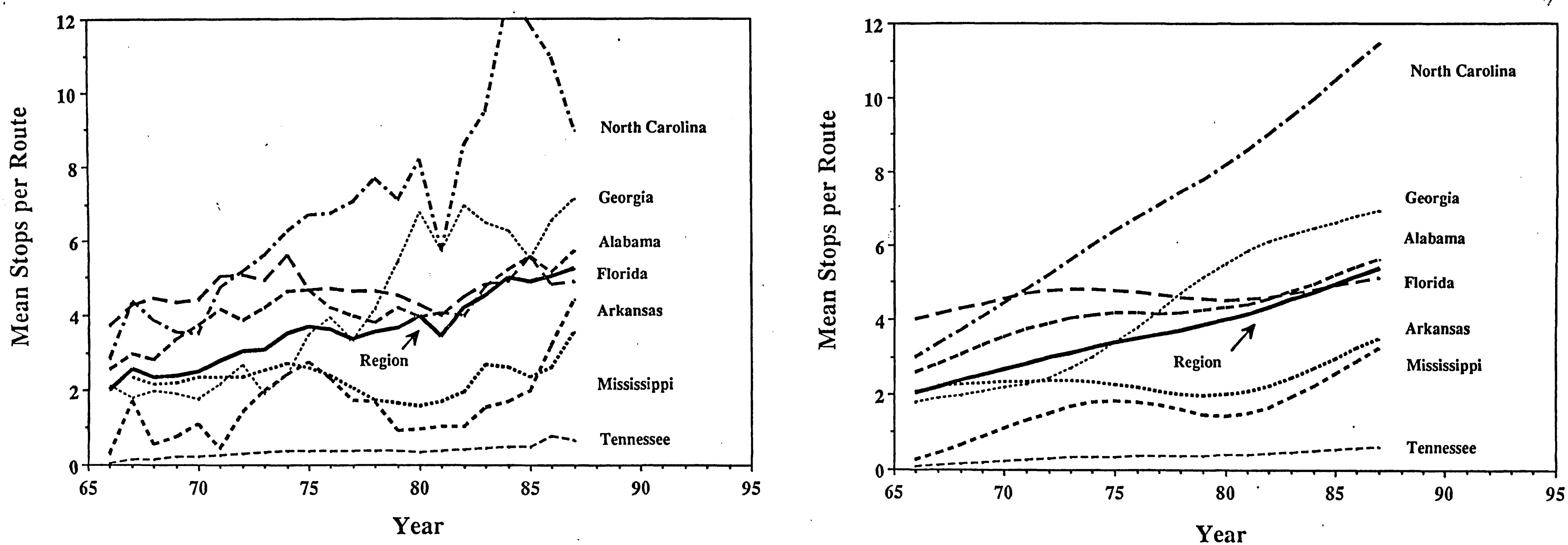
c
Stops per Route
Year
Smoothed
Stratum
Line
Cold Winters
25
20
15
10
5
0
65
70
75
80
85
90

Fig. 5



Arkansas
a
Mean Stops per Route
State Line
4
19
5
Year
Arkansas
b
Mean Stops per Route
State Line
4
19
5
Year
Florida
c
Mean Stops per Route
4
3
State Line
2
Year
Florida
d
Mean Stops per Route
4
3
State Line
2
Year
North Carolina
e
Mean Stops per Route
11
State Line
3
Year
North Carolina
f
Mean Stops per Route
11
3
State Line
Year

a b

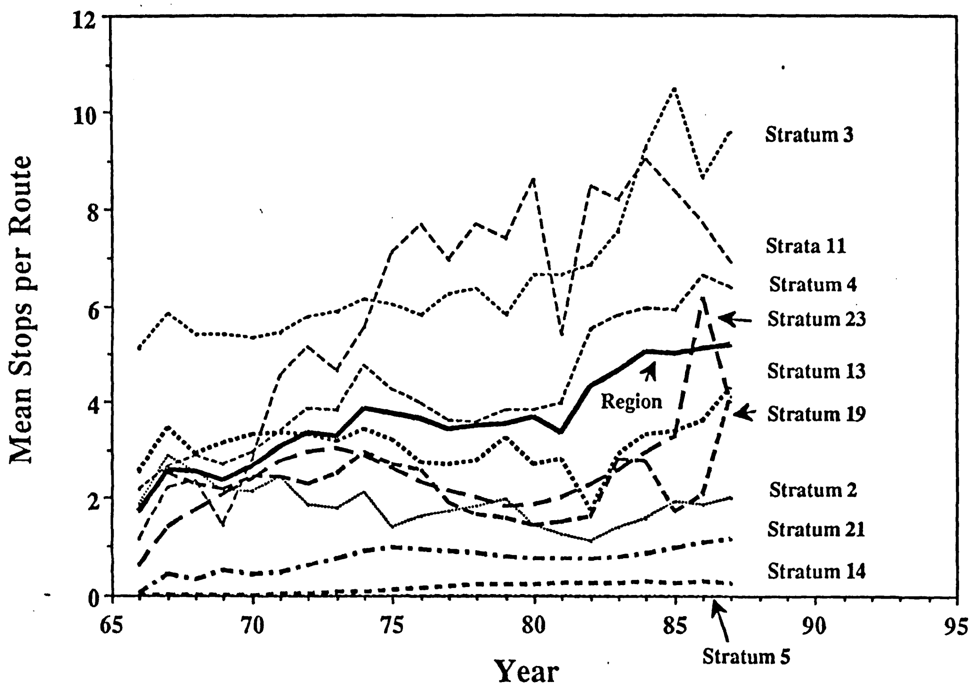


12
10
8
6
4
2
0
65
70
75
80
85
90
95
Mean Stops per Route
Year
North Carolina
Georgia
Alabama
Florida
Arkansas
Mississippi
Tennessee
Region

Fig. 6 a & b

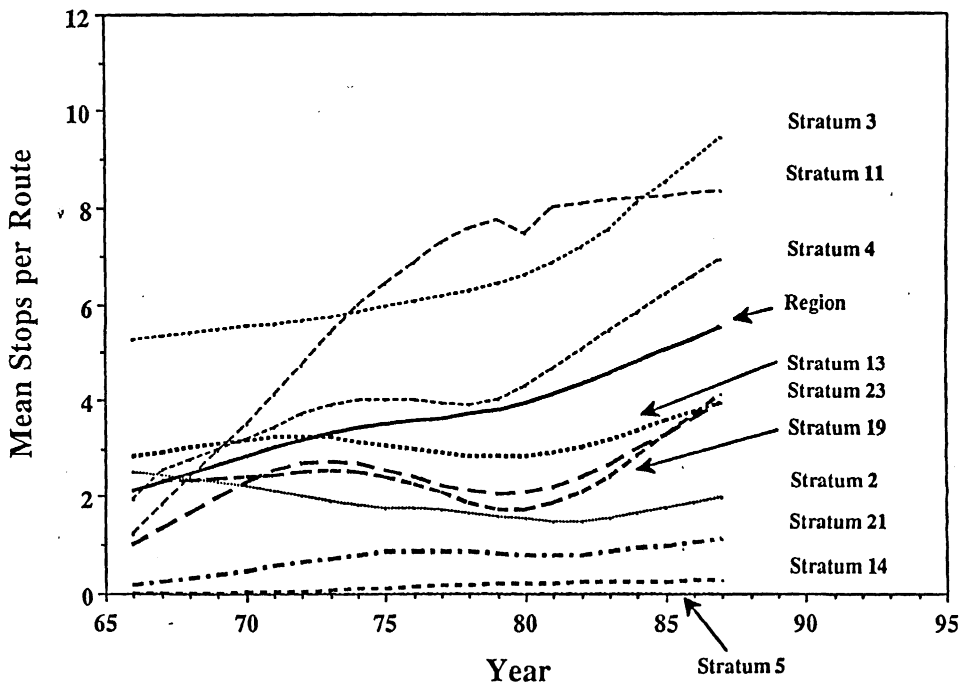
b

c



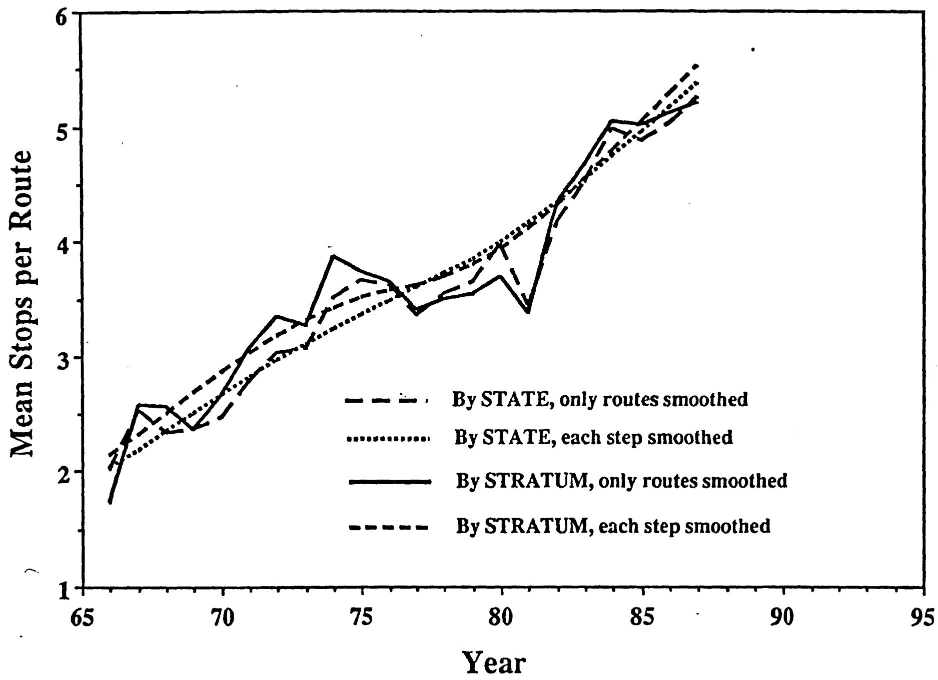
12
10
8
6
4
2
0
65
70
75
80
85
90
95
Mean Stops per Route
Year
Stratum 3
Strata 11
Stratum 4
Stratum 23
Stratum 13
Stratum 19
Region
Stratum 2
Stratum 21
Stratum 14
Stratum 5

d



12
10
8
6
4
2
0
65
70
75
80
85
90
95
Mean Stops per Route
Year
Stratum 3
Stratum 11
Stratum 4
Region
Stratum 13
Stratum 23
Stratum 19
Stratum 2
Stratum 21
Stratum 14
Stratum 5

Fig. 7



Mean Stops per Route
Year
By STATE, only routes smoothed
By STATE, each step smoothed
By STRATUM, only routes smoothed
By STRATUM, each step smoothed

APPENDIX: DATA SMOOTHING PROCEDURE

Analyses were done on an Apple Macintosh computer with the following software:

Systat v3.1 (Systat, Inc., Evanston, IL) for LOWESS smoothing and some data selection;

StatView 512+ (BrainPower Software, Calabasas, CA) for data importing;

Cricket Graph (Cricket Software, Malvern, PA) for plotting;

Absoft FORTRAN (Absoft Corp., Auburn Hills, MI) for writing utility programs.

1. Determine which routes are eligible for inclusion in the analysis. Criteria are:
   a. Type 1 data for at least 10 years out of the 22-year period;
   b. Type 1 data for at least one year in each 5-6-year period. The periods are:
      1966-1971 (6 years)
      1972-1976 (5 years)
      1977-1981 (5 years)
      1982-1987 (6 years)

2. At this time, Systat does not do LOWESS smoothing on a "BY" variable (as in "by route"), so a separate file must be made for each route. LOWESS is run on each file. The column with the smoothed data is then merged back into the original route file to have access to the column of year values. These merged route files are then converted to text format.

3. A utility program (AVGIT) is run on the smoothed route lines to get stratum-within-state averages. The program asks how many and which route files are to be input. A new file is created that contains averages of all smoothed route values of 1966, then for 1967, . . . , 1987. The program is run for each stratum-within-state in the region.

4. The stratum-within-state files (now with one value for each year) are read into Systat. If additional smoothing is desired, LOWESS is run on each file. They are saved in text format.

5. A utility program (MULTIT) reads all the smoothed stratum-within-state files in a particular state. The program then asks for the proportion each stratum is of the entire state area. A weighted average for each year is then calculated for the state. The program is run for each state in the region.

6. The state files are read into Systat. If additional smoothing is desired, LOWESS is run on each one. These are the final smoothed state lines. These are saved in text format.

7. A utility program (SMULTIT) reads all the smoothed state files and asks for the proportion each is of the entire region. A weighted average for each year is calculated for the region.

8. If additional smoothing is desired, the region line is read into Systat and LOWESS is run on it. This is the final smoothed region line.

9. Similarly, steps 5 through 8 can be used to aggregate by strata instead of by states.